Impact of petrophysical uncertainty on Bayesian hydrogeophysical inversion and model selection

Carlotta Brunetti[a], Niklas Linde[a]

[a] *Applied and Environmental Geophysics Group, Institute of Earth Sciences, University of Lausanne, 1015 Lausanne, Switzerland, Carlotta.Brunetti@unil.ch, Niklas.Linde@unil.ch*

Abstract

Quantitative hydrogeophysical studies rely heavily on petrophysical relationships that link geophysical properties to hydrogeological properties and state variables. Coupled inversion studies are frequently based on the questionable assumption that these relationships are perfect (i.e., no scatter). Using synthetic examples and crosshole ground-penetrating radar (GPR) data from the South Oyster Bacterial Transport Site in Virginia, USA, we investigate the impact of spatially-correlated petrophysical uncertainty on inferred posterior porosity and hydraulic conductivity distributions and on Bayes factors used in Bayesian model selection. Our study shows that accounting for petrophysical uncertainty in the inversion (I) decreases bias of the inferred variance of hydrogeological subsurface properties, (II) provides more realistic uncertainty assessment and (III) reduces the overconfidence in the ability of geophysical data to falsify conceptual hydrogeological models.

Keywords: petrophysical uncertainty, hydrogeophysics, Bayesian model selection, Bayesian inversion, evidence, conceptual model

1. Introduction

A primary goal in hydrogeophysical studies is often to infer quantitative hydrogeological models from geophysical and any available hydrogeological data. Unfortunately, petrophysical relationships describing links between geophysical properties and hydrogeological parameters and state variables are uncertain and the information content of hydrogeophysically-inferred estimates is significantly affected by their predictive power. We distinguish here between three types of uncertainty in petrophysical (also called rock physics) models: (1) *petrophysical model uncertainty* refers to uncertainty about the most appropriate parametric form (e.g., Archie's law, time propagation model, Wyllie's formula), (2) *petrophysical parameter uncertainty* relates to uncertainty about the most appropriate parameter values (e.g., cementation index, saturation exponent), and (3) *petrophysical prediction uncertainty* describes the scatter and bias around the calibrated petrophysical model (e.g., dispersion around predictions based on Archie's law). These three types of uncertainty are clearly not independent of each other. For instance, petrophysical prediction uncertainty is described by the residuals between the actual prediction quantity (e.g., porosity, hydraulic conductivity) and the predictions for a given petrophysical model and parameter values.

To date, most focus in hydrogeophysical inversion has been on petrophysical parameter uncertainty (e.g., Kowalsky et al. (2005); Lochbühler et al. (2014)) with the petrophysical parameter values being inferred (deterministically or probabilistically) as a part of the inversion process. However, ignoring the other two types of uncertainty may lead to biased estimates and unrealistically low uncertainty estimates. For instance, Brunetti et al. (2017) suggest that ignoring petrophysical prediction uncertainty when using Bayesian model selection to discriminate among conceptual hydrogeological models will likely lead to over confidence in the ability of geophysical data to falsify and discriminate between alternative conceptual hydrogeological models (Linde, 2014). Furthermore, it also implies that ad hoc data weighting schemes are needed when jointly in-

verting geophysical and hydrogeological data (e.g., Lochbühler et al. (2013) in which each data type was given an equal weight in the objective function).

One approach to partly circumvent these issues is to avoid the use of explicit petrophysical relationships altogether. For instance, this can be achieved using structural approaches to joint inversion (Haber & Oldenburg, 1997). The cross-gradient method of Gallardo & Meju (2003) is a widely employed approach to penalize structural dissimilarity between any two parameter fields (defined as the cross-product of the spatial gradients of two parameter fields). Hydrogeophysical adaptations and applications of this method can be found in Doetsch et al. (2010); Linde et al. (2006, 2008); Lochbühler et al. (2013). Unfortunately, minimizing the cross-gradient function is an inappropriate approach when both hydrogeological properties and state variables vary (e.g., Doetsch et al. (2010); Linde et al. (2006)). Among a multitude of cluster-based approaches, we highlight the works by Sun & Li (2016, 2017) who develop a multidomain joint clustering inversion method that uses the fuzzy c-means clustering technique to constrain the statistical behaviour of inverted physical property values in the parameter domain. This approach overcomes the problem of determining a priori the appropriate petrophysical model as it is allowed to exhibit different forms in different regions of the model domain. For time-lapse applications, Vasco et al. (2014) circumvent the use of an explicit petrophysical model by relating the time at which a significant change in geophysical data occurs to the time of a saturation and/or pressure change within a reservoir or aquifer. Alternative approaches are presented by Hermans et al. (2016) and Oware et al. (2013). They link geophysical properties to hydrogeological parameters by physically-based regularization operators or direct multivariate statistical models but, unlike other methods, they adopt an explicit petrophysical relationship to create a prior set of subsurface model realizations or training images. This is done to ensure geologically realistic results.

Explicit petrophysical relationships can be integrated in hydrogeophysical inversions using two types of work flows: two-step (or sequential) inversion approaches (Chen et al., 2001; Copty et al., 1993; Doyen, 1988, 2007; Rubin

et al., 1992) and coupled inversion approaches (Hinnell et al., 2010; Kowalsky et al., 2005).

The two-step inversion approach consists of two sequential steps: first, the geophysical properties (e.g., electrical permittivity) are inferred from geophysical data (e.g., first-arrival ground-penetrating radar (GPR) travel times) through deterministic or stochastic inversions; second, petrophysical relationships are used to classify and map the inferred geophysical properties into probability density functions (Mukerji et al., 2001) or deterministic estimates of hydrogeological or reservoir properties. This is achieved by different statistical techniques, such as, co-kriging, discriminant analysis, neural networks and Bayesian classification/estimation. In reservoir geophysics, the two-step inversion approach has been favoured in conjunction with sophisticated statistical rock physics models. For instance, Shahraeeni & Curtis (2011); Shahraeeni et al. (2012) use neural networks to map inferred seismic wave impedances into posterior distributions of porosity, clay content, and water saturation. Grana & Della Rossa (2010); Grana et al. (2012) sample the posterior distribution of reservoir properties using the Monte Carlo method for a given seismic model. They conceptualize petrophysical prediction uncertainty as Gaussian random fields with zero mean and a covariance matrix estimated by comparing predictions with well-log data. In hydrogeophysics, the Bayesian two-step approaches are also used, for instance, by Chen et al. (2001, 2004) to estimate hydraulic conductivity conditioned to GPR velocity, GPR attenuation, and seismic velocity tomograms. In hydrogeophysics, the two-step approach has been criticized as it can lead to inconsistent estimates (apparent mass loss) and spatially-dependent bias (Day-Lewis et al., 2005).

The coupled inversion approach is often formulated within a Bayesian framework in which hydrogeological properties are estimated by inversion of geophysical and, possibly, hydrogeological data. A pioneering work on coupled inversion is Bosch (1999) who develops a formal Bayesian procedure, referred to as lithological tomography or lithological inversion. In this approach, Markov chain Monte Carlo (MCMC) is used to integrate geophysical data, geological concepts

and uncertain petrophysical relationships. The coupled inversion approach is well suited to integrate multiple geophysical datasets and arbitrary petrophysical relationships. Also, when confronted with non-linear physics and non-linear petrophysical relationships, the coupled inversion approach is preferable to a two-step inversion approach (Bosch, 2004). Most hydrogeophysical works based on coupled inversion approaches assume that the petrophysical relationship is perfect with known or unknown parameter values (Chen et al., 2006; Kowalsky et al., 2005; Lochbühler et al., 2015). When petrophysical parameter values are unknown, they are inverted for simultaneously with the hydrogeological properties of interest. Petrophysical prediction uncertainty has received less attention in coupled inversion. In the rare circumstances it is included at all, it is commonly conceptualized with a multivariate Gaussian distribution with known mean and covariance matrix (Bosch, 2004; Bosch et al., 2009; Bosch, 2016; Chen & Dickens, 2009). The petrophysical prediction uncertainty is then typically sampled using the brute force Monte Carlo method by adding random multivariate Gaussian realizations to the petrophysical model outputs at each iteration of the MCMC inversion.

In this study, we address the following research questions using a coupled Bayesian hydrogeophysical inversion approach:

1. How can we efficiently incorporate petrophysical prediction uncertainty in MCMC inversions?
2. What are the consequences of ignoring or making incorrect assumptions on petrophysical prediction uncertainty (including its correlation structure) on inferred posterior distributions of interest?
3. Can we reliably infer a geostatistical model of petrophysical prediction uncertainty within the inversion?
4. What are the impacts of petrophysical uncertainty on Bayesian model selection results?

After introducing the theory and method (Section 2), we start out by exploring the above-mentioned research questions by means of porosity estima-

tion using synthetic crosshole GPR travel time data and an explicit well-known petrophysical relationship with known parameters (Section 3). We then present a field case-study (Section 4) aiming at hydraulic conductivity estimation from GPR travel time and hydraulic conductivity (flowmeter) data measured at the South Oyster Bacterial Transport site in Virginia, USA (Chen et al., 2001; Hubbard et al., 2001; Scheibe et al., 2011). Here, we solely assume to know the parametric form of the petrophysical relationship and we infer for its petrophysical parameters (i.e., the petrophysical parameter uncertainty is considered in addition to petrophysical prediction uncertainty).

2. Theory and method

2.1. Bayesian inference and model selection

We present below a short summary of Bayesian inference and model selection.

Given n measurements, $\widetilde{\mathbf{Y}} = \{\widetilde{y}_1, \ldots, \widetilde{y}_n\}$, and a d-dimensional vector of model parameters, $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_d\}$, Bayes' theorem defines the posterior probability density function (pdf) of the model parameters, $p(\boldsymbol{\theta}|\widetilde{\mathbf{Y}})$, as

$$p(\boldsymbol{\theta}|\widetilde{\mathbf{Y}}) = \frac{p(\boldsymbol{\theta})L(\boldsymbol{\theta}|\widetilde{\mathbf{Y}})}{p(\widetilde{\mathbf{Y}})}. \tag{1}$$

The posterior pdf describes the state of knowledge about the model parameters given the observed data and prior knowledge. The prior pdf, $p(\boldsymbol{\theta})$, quantifies the initial state of knowledge about the model parameters before considering the observed data. We consider a likelihood function, $L(\boldsymbol{\theta}|\widetilde{\mathbf{Y}})$, that is Gaussian in shape by imposing uncorrelated and normally distributed measurement errors with constant standard deviation, $\sigma_{\widetilde{\mathbf{Y}}}$,

$$L(\boldsymbol{\theta}|\widetilde{\mathbf{Y}}) = \left(\sqrt{2\pi\sigma^2_{\widetilde{\mathbf{Y}}}}\right)^{-n} \exp\left[-\frac{1}{2}\sum_{h=1}^{n}\left(\frac{\mathcal{F}_h(\boldsymbol{\theta}) - \widetilde{y}_h}{\sigma_{\widetilde{\mathbf{Y}}}}\right)^2\right]. \tag{2}$$

The larger the likelihood, the lower is the data misfit between the simulated forward responses, $\mathcal{F}(\boldsymbol{\theta})$, and the data, $\widetilde{\mathbf{Y}}$. The evidence, $p(\widetilde{\mathbf{Y}})$, evaluates the support provided by the observed data to a given model parametrization and

prior pdf (conceptual model), η, and it is defined as the (multidimensional) integral of the likelihood function over the prior distribution,

$$p(\widetilde{\mathbf{Y}}|\eta) = \int L(\boldsymbol{\theta}, \eta|\widetilde{\mathbf{Y}}) p(\boldsymbol{\theta}|\eta) d\boldsymbol{\theta}. \tag{3}$$

Computing the evidence is challenging as, in general, the integral in Eq. (3) can not be evaluated analytically and it must be approximated by numerical means.

The evidence is used to calculate Bayes factors and is, thus, the cornerstone of Bayesian model selection (Kass & Raftery, 1995). Bayesian model selection (Jeffreys, 1935, 1939) aims at determining the competing conceptual model that is the most supported by the observed data while honouring the principle of Occam's razor. This implies that if multiple conceptual models fit the data nearly equally well, then the simplest model (e.g., with the least number of unknown parameters or the smallest prior parameter ranges) is favoured over more complex ones (Gull, 1988; Jeffreys, 1939; Jefferys & Berger, 1992; MacKay, 1992). Conceptual models could refer to different spatial parametrizations of the subsurface (e.g., multi-Gaussian fields with isotropy or vertical anisotropy) or alternative petrophysical relationships. Bayes factors are simply the ratio of the evidences of two competing conceptual models, η_1 and η_2. For instance, the Bayes factor of η_1 with respect to η_2, or $B_{(\eta_1,\eta_2)}$, is defined as

$$B_{(\eta_1,\eta_2)} = \frac{p(\widetilde{\mathbf{Y}}|\eta_1)}{p(\widetilde{\mathbf{Y}}|\eta_2)}. \tag{4}$$

Subsurface conceptual models with large Bayes factors are preferred statistically and the conceptual model with the largest evidence is the one that best honours the data on average over the prior pdf. This implies that there is no guarantee that the "correct" conceptual model will be favoured if a simpler model allows for similar degrees of data misfit.

In this work, we perform coupled Bayesian hydrogeophysical inversion based on MCMC sampling (Robert & Casella, 2013) using the DREAM$_{(\mathrm{ZS})}$ algorithm (Laloy & Vrugt, 2012; Vrugt, 2016) to estimate $p(\boldsymbol{\theta}|\widetilde{\mathbf{Y}})$. This multi-chain method creates symmetric model proposals from an historical archive of past

states and automatically tunes the scales and orientation of the proposal distribution on the fly to the target posterior distribution. Each proposal is accepted or rejected based on the Metropolis acceptance ratio (Hastings, 1970; Metropolis et al., 1953). If the proposal is accepted, the chain moves to the new location, otherwise the chain remains at its current location. Acceptance ratios between 15% - 40% usually indicate good performance of the MCMC simulation (Gelman et al., 1996). The convergence to the target posterior distribution is monitored with the analysis of variance by Gelman & Rubin (1992). Approximate convergence is declared when the variance between the different chains is lower than the variance within each single chain (Gilks et al., 1995).

For purposes of Bayesian model selection, we estimate the evidence with the Gaussian mixture importance sampling approach recently developed by Volpi et al. (2017). This approach allows for four different sampling methods: reciprocal importance sampling, importance sampling and bridge sampling with geometric and optimal bridge. Following Brunetti et al. (2017), we rely on importance sampling from a Gaussian mixture model that is fitted to the estimated posterior probability density function.

2.2. MC and MCMC sampling of petrophysical prediction uncertainty

As mentioned in Section 1, in the rare cases when petrophysical prediction uncertainty is included in coupled inversion, it is sampled through the brute force Monte Carlo (MC) method (Hammersley & Handscomb, 1964) while the inference of model parameters of interest is achieved through MCMC. This method draws independent samples from the (multivariate) prior distribution of petrophysical prediction uncertainty and we refer to it as MC-within-MCMC. In Section 3.1, we will demonstrate that the MC-within-MCMC method turns out to be very inefficient because of acceptance rates that are prohibitively low. As an alternative, we make use of the DREAM$_{\text{(ZS)}}$ proposal mechanism (see details in Laloy & Vrugt (2012); Vrugt (2016)) to infer the petrophysical prediction uncertainty together with the other parameters by MCMC (full MCMC). In essence, this implies that petrophysical prediction uncertainty is parameterized

and treated in the same way as the other unknowns that are inferred in the MCMC inversion. Both the MC-within-MCMC and the full MCMC approaches should converge to the same result. An alternative to such explicit treatments of petrophysical prediction uncertainty as "nuisance" parameters is to incorporate their effects in the likelihood function. However, efficient and theoretically-consistent ways to achieve this for non-linear problems remains an open research question (see Section 5.2 in Linde et al. (2017)).

2.3. Petrophysical relationships and geophysical forward model

We consider synthetic test cases for known and theoretically-based petrophysical relationships for which petrophysical prediction uncertainty is comparatively low. For the field study, we consider an unknown, empirically-based and comparatively weak petrophysical relationship. The synthetic example concerns predictions of the porosity field and the field study aims at predicting hydraulic conductivity. These two types of problems were chosen to span typical applications, as well as different strengths and types of petrophysical relationships.

The synthetic examples (Section 3) used in this study rely on the following petrophysical relationship to link GPR velocities, $\mathbf{v}$ [m/s], to porosities, $\mathbf{\Phi}$ [-]:

$$\mathbf{v} = \sqrt{\mathbf{\Phi}^m c^{-2}[\varepsilon_{\mathrm{w}} + (\mathbf{\Phi}^{-m} - 1)\varepsilon_{\mathrm{s}}]}^{-1}, \tag{5}$$

where $\varepsilon_{\mathrm{w}} = 81$ [-] and $c = 3 \cdot 10^8$ [m/s] are the relative permittivity of water and the speed of light in vacuum, respectively. We assume the relative permittivity of the mineral grains, ε_{s} [-], equal to 5 and the cementation index, m [-], equal to 1.5. In order to incorporate the petrophysical prediction uncertainty, Eq. (5) is computed in three steps. The effective relative permittivities, $\boldsymbol{\varepsilon}$, are first found for a given porosity model (Pride, 1994):

$$\text{Step 1} : \boldsymbol{\varepsilon} = \varepsilon_{\mathrm{s}} + \mathbf{\Phi}^m \varepsilon_{\mathrm{w}} - \mathbf{\Phi}^m \varepsilon_{\mathrm{s}}, \tag{6}$$

then the petrophysical prediction errors, $\Delta\mathbf{p}$, describing the residual for each model cell are added

$$\text{Step 2} : \boldsymbol{\varepsilon}' = \boldsymbol{\varepsilon} + \Delta\mathbf{p}, \tag{7}$$

and the corresponding GPR velocities are derived

$$\text{Step 3} : \mathbf{v} = \sqrt{c^{-2}\boldsymbol{\varepsilon}'}^{-1}. \tag{8}$$

In the context of the field study (Section 4) at the South Oyster Bacterial Transport Site, we compare linear and quadratic petrophysical relationships to link the GPR velocities, $\mathbf{v}$ [m/s], to the natural logarithm of the hydraulic conductivities, $\boldsymbol{\mathcal{K}} = \log \mathbf{K}$ [log(m/h)]:

$$\text{Step 1} : \mathbf{v}' = a_0 + a_1\boldsymbol{\mathcal{K}} \tag{9}$$

or

$$\text{Step 1} : \mathbf{v}' = a_0 + a_1\boldsymbol{\mathcal{K}} + a_2\boldsymbol{\mathcal{K}}^2 \tag{10}$$

where a_0, a_1 and a_2 are the polynomial coefficients. We then add $\Delta\mathbf{p}$:

$$\text{Step 2} : \mathbf{v} = \mathbf{v}' + \Delta\mathbf{p}. \tag{11}$$

Chen et al. (2001) and Hubbard et al. (2001) demonstrate at the South Oyster Bacterial Transport Site that the GPR velocities inferred by linear tomographic inversion are correlated to the logarithm of hydraulic conductivities with a correlation coefficient of 0.68. This suggests that the true underlying correlation is equal or stronger than this value. However, we stress that any relationship between GPR velocity and hydraulic conductivity is site-specific and typically weak.

The spatial model domain of interest covers an area of 7.2 m $\times$ 7.2 m below the ground surface. We consider multi-Gaussian models of porosity, hydraulic conductivity and petrophysical prediction uncertainty over a regular 2D grid of size 180 $\times$ 180. We use the non-linear 2D traveltime solver (*time 2d*) of Podvin & Lecomte (1991) to compute first-arrival travel times from velocity fields obtained by applying the petrophysical relationships of Eqs. (5), (8) and (9)-(11) to each porosity or hydraulic conductivity field.

2.4. Model parameterisation

We generally describe the petrophysical prediction uncertainty, $\Delta\mathbf{p}$, the porosity, $\boldsymbol{\Phi}$, and the log-hydraulic conductivity, $\boldsymbol{\mathcal{K}}$, fields as multi-Gaussian ran-

dom fields. The only exception is the illustrative synthetic example of Section 3.1, in which the $\mathbf{\Phi}$ and $\mathbf{\Delta p}$ fields correspond to independent horizontal layers. We parameterise our multi-Gaussian fields using the method by Laloy et al. (2015). This method generates stationary multi-Gaussian fields by employing circulant embedding of the covariance matrix. To decrease the number of unknowns inferred during the inversion process, the dimensionality is reduced by resampling two low-dimensional vectors of standard normal random numbers to the original size of the model using the one-dimensional Fast Fourier Transform interpolation. We refer to Laloy et al. (2015) for more details. In our case, we generate each vector with 50 dimensionality reduction (**DR**) variables (i.e., 100 instead of 32400 unknowns), which substantially decrease the MCMC computational cost. The multi-Gaussian model is described by the Matérn variogram model and associated geostatistical parameters, including the mean and the variance, the integral scale along the major axis of anisotropy, I, the anisotropy angle, φ, the ratio of the integral scales along the minor and major axis of anisotropy, R, and the shape parameter of the Matérn variogram model, ν. We jointly infer the geostatistical parameters and the **DR** variables describing the hydrogeological properties (i.e., porosity or hydraulic conductivity) with the corresponding parameters and variables characterising the petrophysical prediction uncertainty.

3. Synthetic examples

3.1. Toy example: MC-within-MCMC versus full MCMC sampling

Historically (see Section 2.2), petrophysical prediction uncertainty has been addressed by drawing independent proposals of $\mathbf{\Delta p}$ from the prior while parameters of interest have been inferred by MCMC (MC-within-MCMC). As an alternative, petrophysical prediction uncertainty is here parameterized and inferred as any other parameter in the MCMC inversion (full MCMC). We consider a toy example to demonstrate the advantage of using an appropriate model proposal distribution to infer the petrophysical prediction uncertainty

(full MCMC) when considering moderately large or large data sets with high signal-to-noise-ratios. The set-up of this simple synthetic example consists of 10 GPR transmitters and 10 receivers placed at uniform depth intervals on the right and left side of the model domain, respectively (Fig. 1a). Considering all possible transmitter-receiver pairs yields 100 first-arrival travel time data. The true porosity field is characterized by four layers of equal thickness with values of 0.3, 0.45, 0.35 and 0.4 starting from the ground surface (Fig. 1a). We consider synthetic travel time data that are contaminated with uncorrelated and normally distributed measurement errors with standard deviation, $\sigma_{\widetilde{\mathbf{Y}}}$, equal to 0.5 ns (i.e., typical of crosshole GPR) and 2 ns, respectively. We consider a uniform prior distribution of porosity in the range [0.25,0.50] and the prior distribution of the petrophysical prediction uncertainty, $\mathbf{\Delta p}$, is Gaussian with zero-mean and standard deviation of 0.8, chosen according to the experimental study of Roth et al. (1990). The $\mathbf{\Delta p}$ values are added following Eq. 7 and integrated in the inversion with the MC-within-MCMC and the full MCMC methods (see Section 2.2). The latter draws the parameters from the DREAM$_{(\mathrm{ZS})}$ proposal distribution that gradually update $\mathbf{\Delta p}$.

We obtain appropriate acceptance rates of 20% (with $\sigma_{\widetilde{\mathbf{Y}}} = 0.5$ ns) and 22% (with $\sigma_{\widetilde{\mathbf{Y}}} = 2.0$ ns) when considering full MCMC (Table 1). For MC-within-MCMC, the acceptance ratio is 0.002% when $\sigma_{\widetilde{\mathbf{Y}}} = 0.5$ ns and 0.31% when $\sigma_{\widetilde{\mathbf{Y}}} = 2.0$ ns. Convergence to the target distribution is consequently much faster for full MCMC than for MC-within-MCMC, especially when $\sigma_{\widetilde{\mathbf{Y}}} = 0.5$ ns (i.e., $5 \cdot 10^3$ forward simulations needed instead of $9.5 \cdot 10^6$, Table 1). That is, the MCMC-derived method allows for an almost 2000-fold decrease in sampling time with respect to the MC-within-MCMC method. This ratio grows further when using smaller $\sigma_{\widetilde{\mathbf{Y}}}$ and more data.

For the case of $\sigma_{\widetilde{\mathbf{Y}}} = 0.5$ ns, we compare the posterior mean porosity fields and associated standard deviations obtained when ignoring $\mathbf{\Delta p}$ (Fig. 1b and 1e), when using the full MCMC (Fig. 1c and 1f) and the MC-within-MCMC estimated $\mathbf{\Delta p}$ (Fig. 1d and 1g). The posterior mean porosity fields obtained in the three cases (Fig. 1b-d) are very similar and agree very well with the

Table 1: First column, method used to sample $\Delta\mathbf{p}$; second column, standard deviation of the measurement errors used to contaminate the data; third column, average acceptance rate; fourth column, number of iterations needed to reach convergence.

Method	$\sigma_{\widetilde{\mathbf{Y}}}$ **[ns]**	**AR [%]**	**T [-]**
Full MCMC	0.5	20.1	$5.0 \cdot 10^3$
	2.0	21.9	$4.0 \cdot 10^3$
MC-within-MCMC	0.5	0.002	$9.5 \cdot 10^6$
	2.0	0.31	$9.6 \cdot 10^4$

true porosity field shown in Fig. 1a. The incorporation of the petrophysical prediction uncertainty results in a standard deviation (Fig. 1f-g) that is ten times higher than for the case without petrophysical prediction uncertainty (Fig. 1e). These results suggest that petrophysical prediction uncertainty has a strong effect on the inferred model uncertainty and that the full MCMC approach is much more efficient than MC-within-MCMC. In the following, we will only present results obtained by the full MCMC approach and recommend it over MC-within-MCMC.

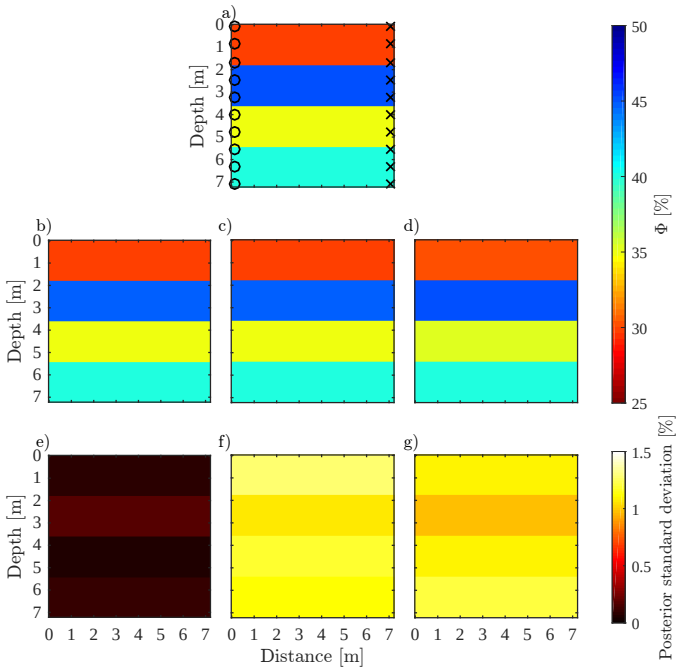


Figure 1: (a) The "true" subsurface porosity model used in our toy example with the different measurement depths of the GPR transmitters (black crosses) and receivers (black circles) indicated. Mean porosity fields of the posterior distribution derived from MCMC simulation with the DREAM$_{(\mathrm{ZS})}$ algorithm using a conceptual model with four layers in the case where (b) the petrophysical prediction uncertainty is not taken into account, (c) the petrophysical prediction uncertainty is sampled by MCMC and (d) the petrophysical prediction uncertainty is sampled by MC-within-MCMC. The corresponding posterior standard deviations of the porosity estimates are shown in (e), (f) and (g), respectively. All these plots were obtained with $\sigma_{\tilde{\mathbf{Y}}} = 0.5$ ns

.

3.2. The forward problem: impact of petrophysical prediction uncertainty

For a given study area, geological facies and properties change in space (e.g., porosity, specific surface area, tortuosity) such that the optimal parameters describing any petrophysical relationship are likely to vary in space. This implies that, when relying on the common assumption of a stationary petrophysical relationship (i.e., the parameter values are the same everywhere), the petro-

physical prediction uncertainty is likely to have a spatially-correlated structure at a scale similar to the geological variability.

In this section, we investigate the impact of spatially-correlated petrophysical prediction uncertainty on data residuals by considering forward responses obtained with and without spatially-correlated petrophysical errors. In this section, we do not perform any inversion, but simply demonstrate the impact of the correlation scale of petrophysical prediction uncertainty. We consider 841 synthetic crosshole GPR travel times that are related to the porosity field in Fig. 2a. The porosity field is described by a multi-Gaussian field with horizontal anisotropy with: $\varphi = 90°$, mean, $\overline{\Phi} = 0.39$, variance, $\sigma^2_{\mathbf{\Phi}} = 2 \cdot 10^{-4}$, integral scale, $I_{\mathbf{\Phi}} = 1.5$ m, integral scales ratio, $R_{\mathbf{\Phi}} = 0.13$ and the shape parameter, $\nu_{\mathbf{\Phi}} = 0.5$ that corresponds to an exponential variogram. In the absence of any petrophysical prediction uncertainty, we obtain the velocity field by applying Eq. 5 with known petrophysical parameters. After calculating the corresponding forward response (Section 2.3), we add uncorrelated Gaussian observational noise with $\sigma_{\widetilde{\mathbf{Y}}} = 0.5$ ns, which leads to a root mean square error (RMSE) of 0.5 ns. For the case of uncorrelated petrophysical prediction errors, we apply Eq. (6), (7) and (8) and draw $\mathbf{\Delta p}$ realizations from an uncorrelated Gaussian distribution with $\sigma_{\mathbf{\Delta p}} = 0.8$. On the resulting simulated travel time data, we add the same observational noise realization. This yields a RMSE of 0.64 ns (Fig. 2b); a comparatively small increase in RMSE compared with the previous case. We then describe the petrophysical prediction uncertainty with zero-mean isotropic ($R_{\mathbf{\Delta p}} = 1$) multi-Gaussian models with $\sigma_{\mathbf{\Delta p}} = 0.8$ and $\nu_{\mathbf{\Delta p}} = 0.5$. To assess the impact of the spatial correlation of the petrophysical prediction uncertainty, we draw $\mathbf{\Delta p}$ realizations for isotropic multi-Gaussian distributions with increasing integral scales. For the corresponding forward responses, we observe a sharp increase of RMSE with increasing integral scales (Fig. 2b). For example, it is higher than 1.20 ns for an integral scale of 1.5 m. The RMSE reaches a plateau slightly above 1.36 ns when the integral scale approaches the size of the model domain (7.2 m). These results suggests that uncorrelated petrophysical prediction uncertainty (i.e., described by a nugget model) will have a relatively weak

impact on inversion results when considering finely-discretized models. However, we suspect petrophysical prediction uncertainty to be spatially-correlated and this correlation increase the effect on the observed data. If these effects are ignored in the inversion, one would expect negative impacts on the inversion results. This is studied in the following section.

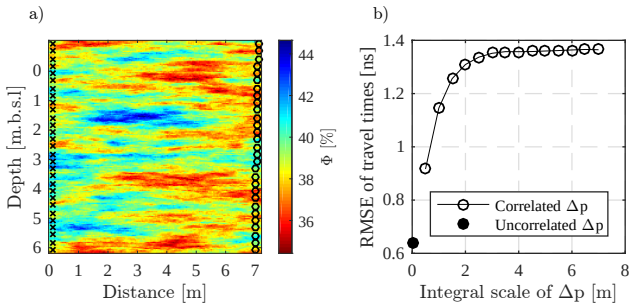


Figure 2: (a) The true porosity model used in our synthetic examples. The 29 GPR transmitter (black crosses) and 29 receiver (black circles) locations are indicated. (b) Root mean square error (RMSE) of GPR travel time data as a consequence of observational errors and petrophysical prediction uncertainty with increasing correlation. In the absence of petrophysical prediction uncertainty, the RMSE is 0.5 ns.

3.3. The inverse problem: impact of assumptions on petrophysical prediction uncertainty

In this section, we investigate the consequences of making incorrect assumptions about petrophysical prediction uncertainty when inferring posterior distributions and Bayesian model selection. We consider the same "true" porosity field (Fig. 2a) as in Section 3.2 and 841 first-arrival GPR travel time data contaminated with uncorrelated and normally-distributed measurement errors with standard deviation, $\sigma_{\tilde{\mathbf{Y}}} = 0.5$ ns. In the MCMC inversions, we infer multi-Gaussian porosity fields with horizontal anisotropy and $\mathbf{DR}_{\mathbf{\Phi}}$, $\overline{\Phi}$, $\sigma^2_{\mathbf{\Phi}}$ being "unknown" parameters drawn from the associated prior distributions listed in Table 2, while all the other geostatistical parameters affecting the porosity structure are kept fixed. The petrophysical prediction uncertainty (if considered) is described as a zero-mean multi-Gaussian field with horizontal anisotropy and

known geostatistical parameters (i.e., only $\mathbf{DR}_{\mathbf{\Delta p}}$ variables are inferred in the inversion, see Table 2). As before, the standard deviation, $\sigma_{\mathbf{\Delta p}}$, was set equal to 0.8 according to the experimental study of Roth et al. (1990). The addition of $\mathbf{DR}_{\mathbf{\Delta p}}$ leads to a decrease in the magnitude of the correlation coefficient (from -1 to -0.81) between the "true" porosity and the "true" GPR velocity values.

Table 2: Geostatistical parameters of the multi-Gaussian models subject to inference (first column), their respective units (second column), range (third column), prior distribution (fourth column), and number (last column). Dimensionality reduction variables, $\mathbf{DR}_{\mathbf{\Phi}}$, mean, $\overline{\Phi}$, and variance, $\sigma^2_{\mathbf{\Phi}}$, of the porosity field; dimensionality reduction variables, $\mathbf{DR}_{\mathbf{\Delta p}}$, of the petrophysical prediction errors.

Parameter	**Units**	**Prior range**	**Prior**	**No.**
$\mathbf{DR}_{\mathbf{\Phi}}$	-	-	Normal	100
$\overline{\Phi}$	-	$[0.3, 0.5]$	Uniform	1
$\sigma^2_{\mathbf{\Phi}}$	-	$[10^{-4}, 2.5 \cdot 10^{-3}]$	Log-uniform	1
$\mathbf{DR}_{\mathbf{\Delta p}}$	-	-	Normal	100

We consider four cases: $\mathbf{\Delta p}$ is not present in the data (i.e., it is not used to generate the synthetic data) and it is not inferred in the MCMC inversion (Case 1); $\mathbf{\Delta p}$ is inferred but it is not present in the data used for inversion (Case 2); $\mathbf{\Delta p}$ is present in the data, but not inferred (Case 3); $\mathbf{\Delta p}$ is present in the data and inferred (Case 4). Cases 1 and 4 represent situations where the assumptions are consistent with the "field" situation, while Cases 2 and 3 are based on inconsistent assumptions. We suggest that Case 3 represent the most common situation in the hydrogeophysics literature (i.e., petrophysical prediction uncertainty exists, but it is ignored).

All cases considered provide accurate estimates of the mean porosity (Fig. 3a), but only the consistent cases (Case 1 and 4) give significant probability to the actual variance (i.e., sill) describing the porosity field (Fig. 3b), with (as expected) Case 4 providing less precise estimates (i.e., parameter uncertainty is higher). For the inconsistent cases, we find for Case 2 that the standard deviation of the porosity field is greatly underestimated, while it is overestimated

in Case 3 (Fig. 3b).

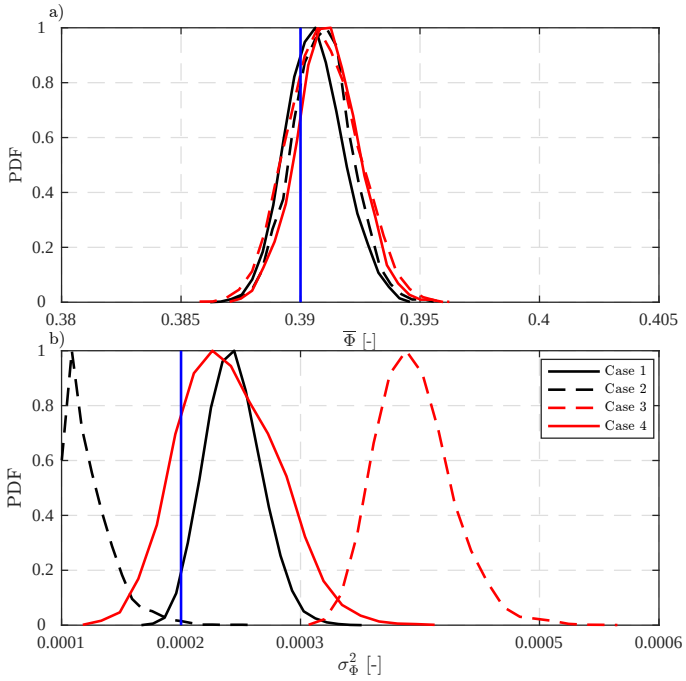


Figure 3: (a) Posterior distributions of the inferred mean of the porosity field. (b) Posterior distributions of the inferred variance (i.e., sill) of the porosity field. The vertical blue lines depict the values of the true model. The posterior distributions are derived from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations.

We now consider the resulting mean porosity fields and the standard deviations for the consistent cases. For Case 1, we find a mean porosity field (Fig. 4a) that is very close to the true field (Fig. 2a). The standard deviation is low (Fig. 4e), the scatter between the mean model and the true model follows the 1:1 trend line (Fig. 4i) and the correlation coefficient is high (0.9). For Case 4, we find a slightly less precise mean model (Fig. 4d), which is reflected in the standard deviation being twice as large (Fig. 4h). Nevertheless, the corresponding scatter plot (Fig. 4l) indicates that there is no bias (the scatter falls on the 1:1 trend line) and the correlation coefficient is 0.75.

We now turn our attention to the inconsistent cases. When considering Case 2, we find a less variable mean field (Fig. 4b) and standard deviations that are

in-between the two consistent cases (Fig. 4f). The correlation coefficient is high (0.88), but the estimates are biased as they do not follow the 1:1 trend line (Fig. 4j). For Case 3, we find an overly variable mean field (Fig. 4c), rather small standard deviations (Fig. 4g) and a moderate correlation coefficient (0.75) with a scatter plot above the 1:1 trend line (Fig. 4k). These results suggest different outcomes. First, including a known petrophysical prediction uncertainty in the inversion leads to consistent estimates, but a wider posterior distribution than if petrophysical prediction uncertainty is absent. Second, the correlation coefficient with the true model is mainly determined by the petrophysical prediction uncertainty. Third, the estimated petrophysical prediction uncertainty (that does not exist) in Case 2 accounts for some of the variability due to porosity variations, which leads to a too smooth mean porosity field. Lastly, ignoring actual petrophysical prediction uncertainty in the inversion process (Case 3; the common case) leads to overly variable fields in order to accommodate data variability caused by both porosity variations and petrophysical prediction uncertainty. From these first inversion examples, we conclude that ignoring petrophysical prediction uncertainty leads to overly confident parameter inference and that some of the estimated parameters might be biased.

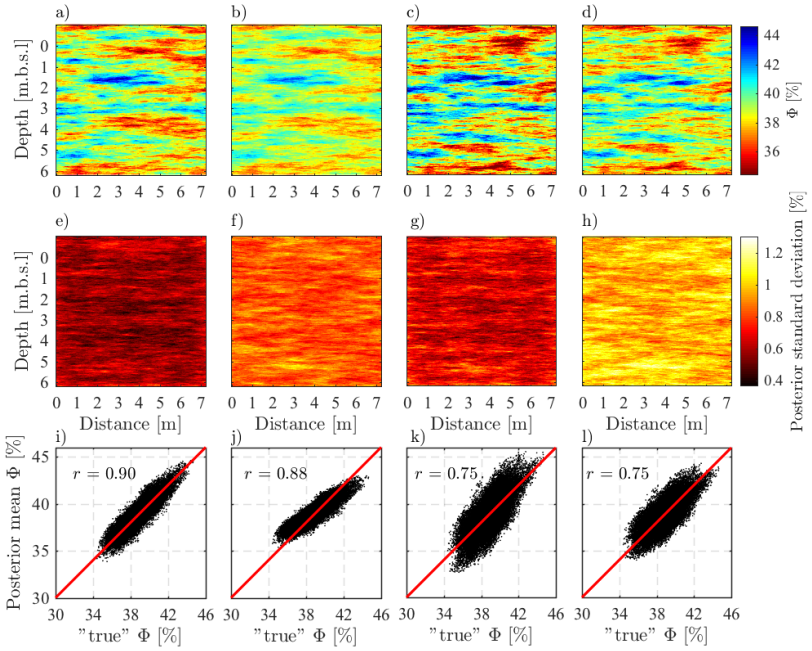


Figure 4: (a-d) Mean porosity fields of the posterior distribution derived from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations for Cases 1-4, respectively. The corresponding posterior standard deviations of the porosity estimates for the four different cases are shown in (e-h), respectively. From (j) to (l), scatter plots of the "true" porosity values versus the mean posterior porosity estimates obtained in the four cases. In each plot, from (j) to (l), the Pearson correlation coefficients, r, are reported and the red lines depict the theoretical 1:1 trend line (i.e., Pearson correlation coefficient equal to 1).

We now focus our attention on Bayesian model selection. For each of the four cases, we also use the data to infer porosity fields assuming (erroneously) a multi-Gaussian conceptual model with isotropy or vertical anisotropy. We compute the evidence for each of these conceptual models (the case of the true horizontal anisotropy and the incorrect cases of isotropy and vertical anisotropy) by approximating the integral in Eq. (3) with the Gaussian mixture importance sampling estimator (Section 2.1). For each case, we use a total of 10^5 importance samples and repeat the evidence computation 10 times. The mean evidences

and associated ranges are presented in Fig. 5.

We find that the ranking of the different conceptual models is the same for all cases. As expected, the multi-Gaussian model with horizontal anisotropy (true conceptual model) has the largest evidence followed by the isotropic model (Fig. 5a). The evidence values are the largest when no petrophysical prediction uncertainty is present in the data or in the inversion (Case 1, Fig. 5a). When we include $\mathbf{\Delta p}$ in the inversion, the evidence estimates (Case 2, Fig. 5a) decrease drastically with respect to Case 1. For instance, we find a 29 orders of magnitude decrease of the evidence estimates for the best model (multi-Gaussian model with horizontal anisotropy). When petrophysical prediction uncertainty is absent in the data (Cases 1 and 2), we find thus that Bayesian model selection clearly indicates that the conceptual model with horizontal anisotropy and no petrophysical prediction uncertainty is superior (the consistent case). Note that this is the case despite the fact that we find the highest log-likelihoods for Case 2 (black dotted lines in Fig. 5b-d). The addition of 100 "unnecessary" degrees of freedom in Case 2 leads to a much decreased ability to differentiate among the different geostatistical models. The error bars of the evidence estimates overlap for Case 2 and the Bayes factors (Table 3) are much smaller than for Case 1, which imply that it is much more difficult to judge which geostatistical model is preferred statistically.

We have seen above that the Bayesian model selection clearly favours the consistent Case 1 when comparing Cases 1 and 2. Unfortunately, this is not the case when comparing Cases 3 and 4. The consistent Case 4 (petrophysical prediction error in data and model parameterization) has a much lower evidence (Fig. 5a) for the multi-Gaussian model with horizontal anisotropy and much lower Bayes factors (Table 3) than the inconsistent Case 3 (petrophysical prediction errors in the data only). The reason for this is that Case 3 has similar log-likelihoods (i.e., data misfit) as Case 4 (Fig. 5b), but half as many model parameters. The ability to fit the data so well with this inconsistent model is probably a consequence of the petrophysical prediction uncertainty having the same geostatistical model as the porosity field. This implies that formal

Bayesian model selection will favour a lower-dimensional model parameterization that fits the data well, regardless of if it is the "correct" model or not. This is a characteristic of Bayesian model selection (e.g., Schöniger et al. (2015b)). Additional tests were performed (not shown) with conditioning to 17 porosity values along each borehole. This decreased the evidence for Case 3 somewhat and increased it for Case 4. However, Case 3 was still strongly favoured when calculating the corresponding Bayes factor.

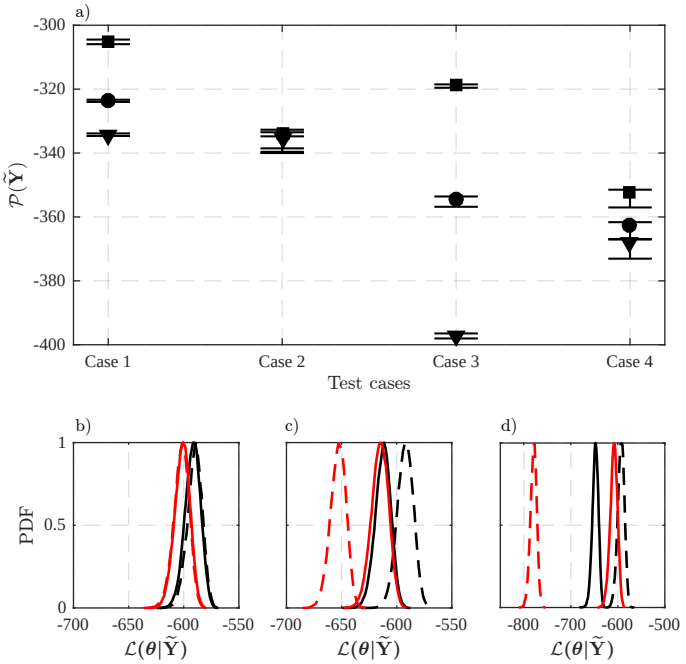


Figure 5: (a) Mean values of the evidence in $\log_{10}$ space, $\mathcal{P}(\widetilde{\mathbf{Y}})$, and corresponding uncertainty (error bars) derived from the Gaussian mixture importance sampling method for the multi-Gaussian conceptual models with horizontal anisotropy (squares), isotropy (circles) and vertical anisotropy (triangles). Posterior distribution of the log-likelihood, $\mathcal{L}(\boldsymbol{\theta}|\widetilde{\mathbf{Y}})$, for the multi-Gaussian model with (b) horizontal anisotropy, (c) isotropy and (d) vertical anisotropy in Case 1 (black solid line), Case 2 (black dotted line), Case 3 (red dotted line) and Case 4 (red solid line).

Table 3: Bayes factors in $\log_{10}$ space of the best conceptual model, MGha, (horizontal anisotropy) with respect to the isotropic one, MGis, (first column) and to the vertically anisotropic one, MGva (last column).

Cases	$\log_{10} B_{(\text{MGha,MGis})}$	$\log_{10} B_{(\text{MGha,MGva})}$
Case 1	18.36	29.09
Case 2	0.58	1.94
Case 3	35.65	78.38
Case 4	10.19	15.37

3.4. Inference of petrophysical prediction uncertainty

We have shown (Section 3.3) that ignoring petrophysical prediction uncertainty in MCMC inversions leads to over confident parameter estimates and biased estimates of geostatistical properties (e.g., the sill). In practical field situations, it is difficult to determine a priori the appropriate geostatistical model that governs petrophysical prediction uncertainty. In this section, we explore to which extent it is possible to infer for both $\mathbf{\Delta p}$ and its underlying geostatistical model. We consider the same overall setting as in Sections 3.2 and 3.3 and the same "true" porosity field (Fig. 2a). Here, the true petrophysical prediction uncertainty is a zero-mean isotropic multi-Gaussian field with $\sigma_{\mathbf{\Delta p}} = 0.8$, $I_{\mathbf{\Delta p}}$=0.8 m, $R_{\mathbf{\Delta p}}$=1, and $\nu_{\mathbf{\Delta p}}$=0.5. We then infer for the mean and variance of the porosity field and for all the geostatistical parameters of $\mathbf{\Delta p}$ described above and the corresponding $\mathbf{DR}_{\mathbf{\Delta p}}$ variables. The corresponding prior distributions of these "unknown" parameters are listed in Tables 2 and 4. The petrophysical relationship used is Eq. (5) and the petrophysical prediction uncertainty is accounted for following Eq. (7).

The inferred posterior distributions of the mean (Fig. 6a) and variance (Fig. 6b) of the porosity field are in general quite well recovered, even if they show a slight tendency to underestimate the true values. Overall, the geostatistical properties of the reference petrophysical prediction uncertainty field are captured in the sense that the corresponding "true" values are included in the

Table 4: Geostatistical parameters of the multi-Gaussian model of the petrophysical prediction uncertainty subject to inference (first column), their respective units (second column), range (third column), prior distribution (fourth column), and number (last column). Standard deviation, $\sigma_{\mathbf{\Delta p}}$, integral scale along the major axis of anisotropy, $I_{\mathbf{\Delta p}}$, anisotropy angle, $\varphi_{\mathbf{\Delta p}}$, ratio of the integral scales, $R_{\mathbf{\Delta p}}$, and shape parameter of the Matérn variogram, $\nu_{\mathbf{\Delta p}}$, of the petrophysical prediction uncertainty field.

Parameter	**Units**	**Prior range**	**Prior**	**No.**
$\sigma_{\mathbf{\Delta p}}$	-	$[0.2, 3.6]$	Log-uniform	1
$I_{\mathbf{\Delta p}}$	m	$[0.6, 3]$	Uniform	1
$\varphi_{\mathbf{\Delta p}}$	$^\circ$	$[0, 180]$	Uniform	1
$R_{\mathbf{\Delta p}}$	-	$[0.05, 1]$	Uniform	1
$\nu_{\mathbf{\Delta p}}$	-	$[0.1, 5]$	Log-uniform	1

posterior distributions (Fig. 6c-g). However, some of the parameters are poorly recovered. For instance, the inferred standard deviation of $\mathbf{\Delta p}$ is centered on the value of 1 instead of 0.8 (Fig. 6c) and the inferred shape parameter of the Matérn variogram peaks on a value that is half of the corresponding "true" value (Fig. 6g). The anisotropy angle is poorly estimated, which is a consequence of the "true" $\mathbf{\Delta p}$ field being isotropic (Fig. 6e). The integral scale along the major axis of anisotropy and the ratio of the integral scales peak on the "true" values, but their posterior distributions are relatively wide (Fig. 6d and 6f).

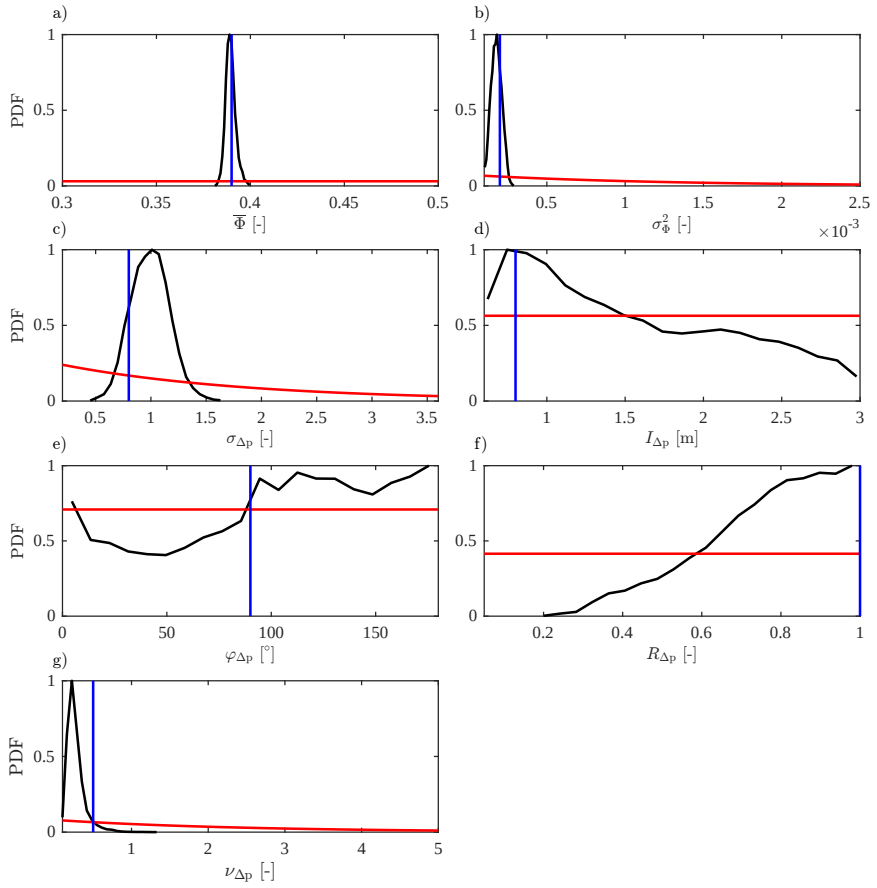


Figure 6: Posterior distributions (black lines) derived from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations of the (a) inferred mean, $\overline{\Phi}$, and (b) variance, σ^2_Φ, of the porosity field and of the geostatistical parameters of the petrophysical prediction uncertainty field: (c) standard deviation, $\sigma_{\mathbf{\Delta p}}$, (d) integral scale along the major axis of anisotropy, $I_{\mathbf{\Delta p}}$, (e) anisotropy angle, $\varphi_{\mathbf{\Delta p}}$, (f) ratio of the integral scales along the minor and major axis of anisotropy, $R_{\mathbf{\Delta p}}$, and (g) shape parameter of the Matérn variogram, $\nu_{\mathbf{\Delta p}}$. The red and blue lines depict the corresponding prior distributions and values of the reference field, respectively. The densities in each plot are normalized.

The dominant structures in the reference porosity field (Fig. 7a), such as the low-porosity zones at a depth of 0.5 m, 4 m and 6 m, are well represented by the posterior mean porosity field (Fig. 7b). The posterior standard deviations

on the inferred porosity field span a range between 0.6% and 1% (Fig. 7c). We find that the inferred mean petrophysical prediction uncertainty field (Fig. 7d) and the "true" field (Fig. 7e) have a rather low correlation coefficient (0.55). The posterior standard deviations of $\Delta\mathbf{p}$ span a range between 0.6 and 1 (Fig. 7f). These large uncertainties are also reflected in the $\Delta\mathbf{p}$ posterior realizations (Fig. 8) that appear to be rather isotropic but with integral scales that vary significantly. Overall, the structural features of the GPR velocity field are well inferred even if their values span a wider range than the reference field (Fig. 7g-h). In particular, the high-velocity zone in the bottom right corner of the model domain are enhanced and characterized by large uncertainties (Fig. 7i).

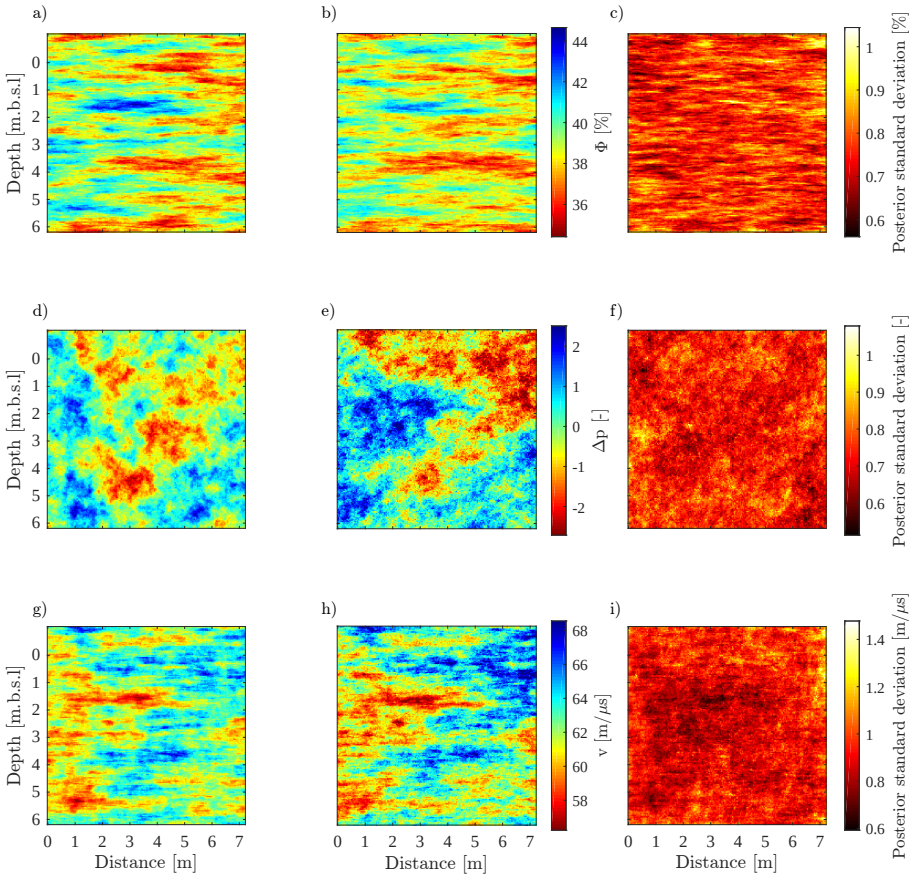


Figure 7: (a) The "true" subsurface porosity model used in our synthetic example; (b) mean porosity field of the posterior distribution derived from MCMC simulation and the corresponding (c) standard deviations. (d) The "true" petrophysical prediction uncertainty model; (e) mean petrophysical prediction uncertainty field of the posterior distribution derived from MCMC simulation and the corresponding (f) standard deviations. (g) The "true" GPR velocity model; (h) mean velocity field of the posterior distribution derived from MCMC simulation and the corresponding (i) standard deviations. The mean fields are obtained from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations.

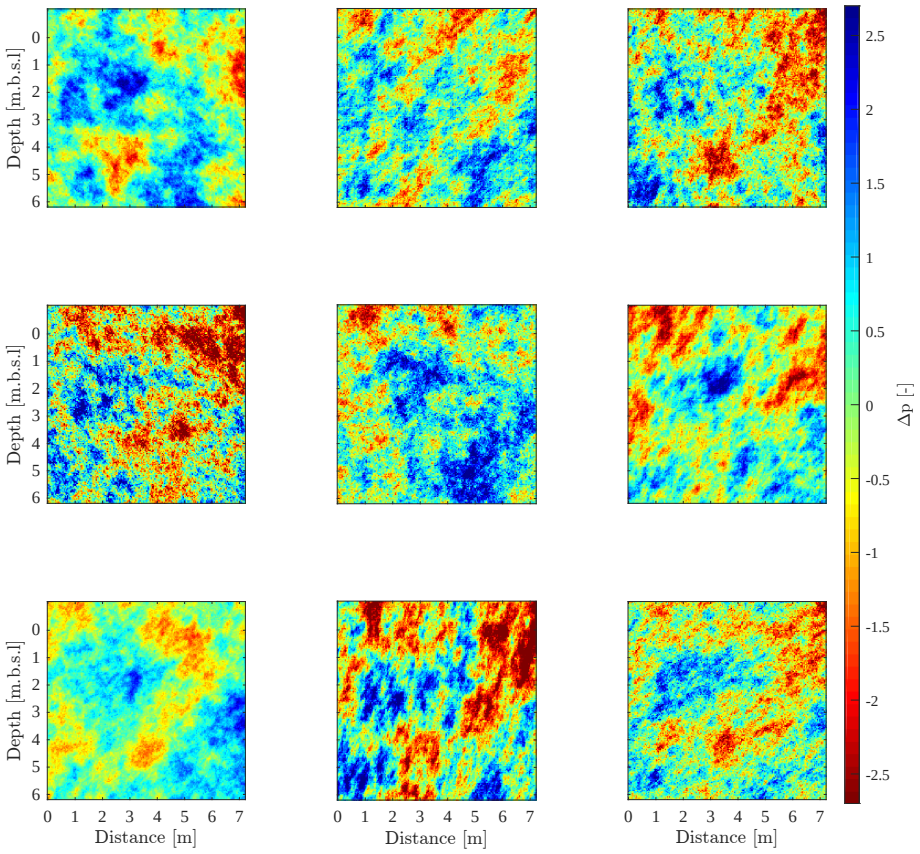


Figure 8: Nine realizations of the petrophysical prediction uncertainty field drawn randomly from the posterior distribution obtained from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations. The petrophysical prediction uncertainty is conceptualized by a multi-Gaussian field with isotropy.

We performed also a test with the petrophysical prediction uncertainty field conceptualized by a multi-Gaussian field with anisotropy at 45° (not shown). For this case, we find a significant improvement in the ability to infer for the standard deviation, angle of anisotropy and the shape parameter of $\mathbf{\Delta p}$. These results suggest that $\mathbf{\Delta p}$ is best resolved when its geostatistical properties are markedly different from the underlying porosity field. However, Bayesian model selection between the two conceptual models that include and not include $\mathbf{\Delta p}$ in

the inversion still favours the case in which petrophysical prediction uncertainty errors are ignored (not shown).

4. Field example

4.1. Field site and available data

We now focus our attention on field data from the South Oyster Bacterial Transport Site in Virginia, USA (Hubbard et al., 2001). In Section 3, we considered a well known and strong petrophysical relationship, while here we consider a case of an unknown and only moderately strong petrophysical relationship. A PulseEKKO 100 GPR system with a 100-MHz nominal-frequency antenna was used and we consider 841 crosshole GPR first-arrival travel time data between 29 transmitter and 29 receiver locations in boreholes S14 and M3, respectively. A total of 95 hydraulic conductivity estimates along boreholes S14, T2 and M13 obtained from an electromagnetic flowmeter were used for point conditioning following the methodology outlined by Laloy et al. (2015). We use the GPR data to infer the underlying log-hydraulic conductivity field, $\boldsymbol{\mathcal{K}}$, assuming a multi-Gaussian model with horizontal anisotropy. Its integral scales, the anisotropy angle, and the shape parameter of the Matérn variogram are set based on previous investigations at the site (Chen et al., 2001; Hubbard et al., 2001). These fixed parameters include, $I_{\boldsymbol{\mathcal{K}}} = 1.5$ m, $\varphi_{\boldsymbol{\mathcal{K}}} = 90°$, $R_{\boldsymbol{\mathcal{K}}} \approx 0.13$ and $\nu_{\boldsymbol{\mathcal{K}}} = 0.5$. The dimensionality reduction variables, $\mathbf{DR}_{\boldsymbol{\mathcal{K}}}$, the mean, $\overline{\mathcal{K}}$, and standard deviation, $\sigma_{\boldsymbol{\mathcal{K}}}$, of the log-hydraulic conductivity field are subject to inference and the corresponding prior ranges are listed in Table 5. The prior range on $\sigma_{\boldsymbol{\mathcal{K}}}$ is set to include the 0.42 log(m/h) standard deviation of the available flowmeter data. The petrophysical prediction uncertainty is described by a zero-mean multi-Gaussian field with prior distributions outlined in Table 5. The upper bound on the prior range of $\sigma_{\Delta\mathbf{p}}$ is chosen such that the resulting correlation coefficient between GPR velocities and log-hydraulic conductivities is equal or stronger than 0.68, which corresponds to the value reported by Chen et al. (2001) and Hubbard et al. (2001). We also jointly infer the petrophysical

parameters a_0, a_1 and a_2 in Eqs. (9)-(10) and the standard deviation of the measurement errors, $\sigma_{\tilde{\mathbf{Y}}}$ (Table 5). The overall number of parameters subject to inference is 211.

Table 5: Parameters subject to inference at the South Oyster Bacterial Transport Site (first column), their respective units (second column), range (third column), prior distribution (fourth column), and number (last column). Dimensionality reduction variables, $\mathbf{DR}_{\boldsymbol{\mathcal{K}}}$, mean, $\overline{\mathcal{K}}$, and standard deviation, $\sigma_{\boldsymbol{\mathcal{K}}}$, of the natural log-hydraulic conductivity field; dimensionality reduction variables, $\mathbf{DR}_{\boldsymbol{\Delta p}}$, standard deviation, $\sigma_{\boldsymbol{\Delta p}}$, integral scale along the major axis of anisotropy, $I_{\boldsymbol{\Delta p}}$, anisotropy angle, $\varphi_{\boldsymbol{\Delta p}}$, ratio of the integral scales, $R_{\boldsymbol{\Delta p}}$, and shape parameter of the Matérn variogram, $\nu_{\boldsymbol{\Delta p}}$, of the petrophysical prediction uncertainty field; standard deviation of the measurement errors on the travel time data, $\sigma_{\tilde{\mathbf{Y}}}$, and polynomial coefficients of the constant, a_0, the linear, a_1, and quadratic, a_2, terms used to describe linear or a quadratic petrophysical relationships.

Parameter	**Units**	**Prior range**	**Prior**	**No.**
$\mathbf{DR}_{\boldsymbol{\mathcal{K}}}$	-	-	Normal	100
$\overline{\mathcal{K}}$	log(m/h)	$[-2, -1]$	Uniform	1
$\sigma_{\boldsymbol{\mathcal{K}}}$	log(m/h)	$[0.4, 0.5]$	Log-uniform	1
$\mathbf{DR}_{\boldsymbol{\Delta p}}$	-	-	Normal	100
$\sigma_{\boldsymbol{\Delta p}}$	m/µs	$[0, 0.8]$	Uniform	1
$I_{\boldsymbol{\Delta p}}$	m	$[0.6, 3]$	Uniform	1
$\varphi_{\boldsymbol{\Delta p}}$	°	$[0, 180]$	Uniform	1
$R_{\boldsymbol{\Delta p}}$	-	$[0.05, 1]$	Uniform	1
$\nu_{\boldsymbol{\Delta p}}$	-	$[0.1, 5]$	Log-uniform	1
$\sigma_{\tilde{\mathbf{Y}}}$	ns	$[0.3, 2]$	Log-uniform	1
a_0	m/µs	$[40, 100]$	Uniform	1
a_1	$\log(h/m)\cdot$ m/µs	$[0, 80]$	Uniform	1
a_2	$\log(h^2/m^2)\cdot$ m/µs	$[0, 5]$	Uniform	1

4.2. Results at the South Oyster Bacterial Transport Site

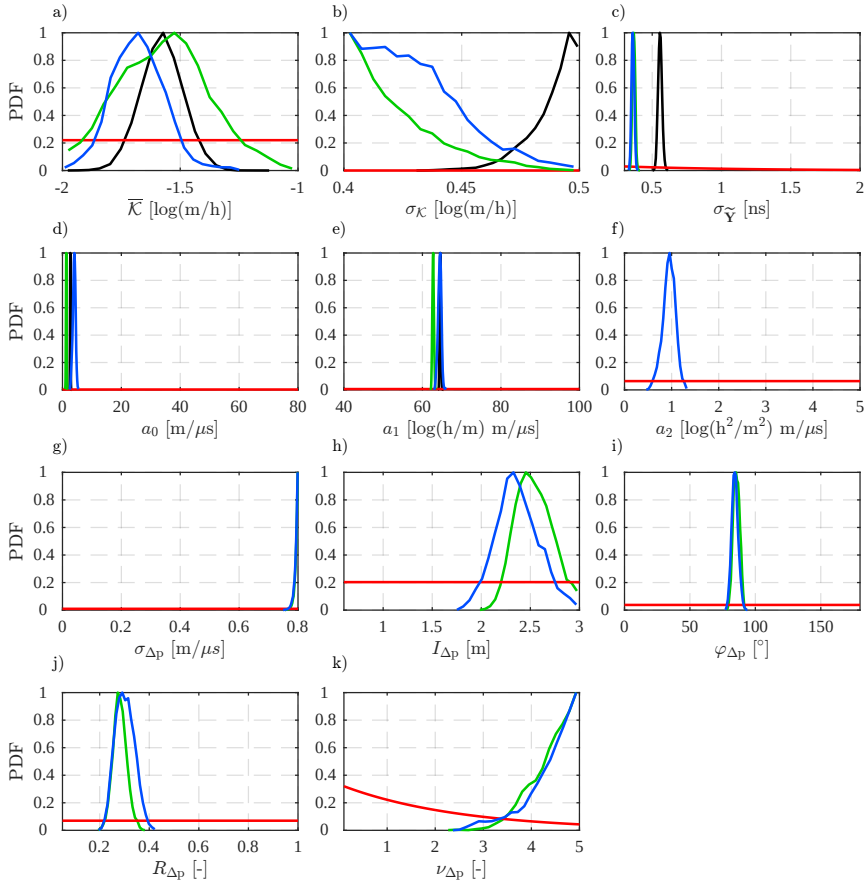
In Section 3, we considered a synthetic example and a known petrophysical relationship. In the present field example, we only assume to know the para-

metric form of the petrophysical relationship and we estimate its petrophysical parameters. We infer the underlying log-hydraulic conductivity field and compare the results obtained by assuming three different petrophysical models: a perfect linear petrophysical relationship (Eq. (9)) in which the petrophysical prediction uncertainty is ignored (Model 1), a linear petrophysical relationship with scatter $\mathbf{\Delta p}$ taken into account by following Eqs. (9) and (11) (Model 2), and a quadratic petrophysical relationship with scatter $\mathbf{\Delta p}$ accounted for as in Eqs. (10)-(11) (Model 3).

After MCMC inversion, we obtain similar posterior distributions of the mean log-hydraulic conductivity when using a perfect linear (-1.58 log(m/h)) and a scattered linear (-1.57 log(m/h)) petrophysical relationship and a slightly lower value (-1.68 log(m/h)) when using a scattered quadratic petrophysical relationship (Fig. 9a). When ignoring $\mathbf{\Delta p}$, the inferred standard deviation of the log-hydraulic conductivity field peaks close to the upper bound (black line, Fig. 9b). When using a scattered linear or quadratic petrophysical relationship, the inferred posterior distribution of the standard deviation is truncated on the lower bound of the prior range (green and blue lines, Fig. 9b). The highest inferred standard deviation of the measurement errors, 0.56 ns, is obtained when ignoring $\mathbf{\Delta p}$ in the inversion (black line, Fig. 9c). When considering the scattered linear or quadratic petrophysical relationship, the corresponding estimates are 0.37 ns and 0.36 ns, respectively (Fig. 9c).

The parameters describing the three petrophysical relationships are well defined (Fig. 9d-e-f). The inferred standard deviation of the petrophysical prediction uncertainty peak on the upper bound of the prior range (Fig. 9g). The other geostatistical parameters describing the $\mathbf{\Delta p}$ field have similar posterior distributions regardless of if a linear (green lines) or a quadratic (blue lines) petrophysical relationship is used (Fig. 9h-k). In particular, we find that the petrophysical prediction uncertainty field is characterized by an integral scale along the major axis of anisotropy centred around 2.4 m (Fig. 9h), an almost horizontal anisotropy (Fig. 9i) and a ratio of the integral scales of 0.30 (Fig. 9j). The posterior distribution of the Matérn shape parameter is truncated by

605 the upper bound, thereby, suggesting a smooth field (Fig. 9k).



606

Figure 9: Posterior distributions derived from MCMC simulation with the DREAM$_{\mathrm{(ZS)}}$ algorithm using 8 chains with $5{\cdot}10^5$ iterations of the (a) mean, $\overline{\mathcal{K}}$, and (b) standard deviation, $\sigma_{\boldsymbol{\mathcal{K}}}$, of the log-hydraulic conductivity field. Posterior distributions of the (c) standard deviation of the measurement errors, $\sigma_{\widetilde{\mathbf{Y}}}$, and the polynomial coefficients of the (d) constant, a_0, (e) linear, a_1 and (f) quadratic, a_2, terms describing the petrophysical relationships of Eqs. (9)-(11). Posterior distributions of the geostatistical parameters of the petrophysical prediction uncertainty field: (g) standard deviation, $\sigma_{\boldsymbol{\Delta p}}$, (h) integral scale along the major axis of anisotropy, $I_{\boldsymbol{\Delta p}}$, (i) anisotropy angle, $\varphi_{\boldsymbol{\Delta p}}$, (j) ratio of the integral scales along the minor and major axis of anisotropy, $R_{\boldsymbol{\Delta p}}$, and (k) shape parameter of the Matérn variogram, $\nu_{\boldsymbol{\Delta p}}$. The results for the perfect linear, scattered linear and scattered quadratic petrophysical relationship are depicted with black, green and blue lines, respectively. The red lines indicate
607 the corresponding prior distributions. The densities in each plot are normalized.

In Fig. 10a-c, we display the mean posterior hydraulic conductivity fields in linear scale. The three fields show similar values close to the boreholes where flowmeter data are available but, away from these locations, the different petrophysical models lead to different subsurface structures and estimates (e.g., within the first meter below the ground surface and between borehole T2 and M3, Fig. 10a-c). Nevertheless, all the three hydraulic conductivity mean models depict a low-hydraulic conductivity zone at a depth of 1-2 m.b.s.l. and at 5-6 m.b.s.l. (Fig. 10a-c). When the petrophysical prediction uncertainty is ignored, the inferred hydraulic conductivity (Fig. 10a) and GPR velocity (Fig. 10g) fields are characterized by a high variability. On average, the standard deviations of the posterior hydraulic conductivity estimates are higher when petrophysical prediction uncertainty is accounted for (Fig. 10d-f).

We observe similarities between the corresponding posterior GPR mean velocities (Fig. 10g-i). For instance, they all show a low-velocity zone within the first 2 m.b.s.l, at 3 m.b.s.l. and at 5-6 m.b.s.l and a high-velocity zone at 4-5 m.b.s.l. As expected, the inferred velocity fields derived from scattered petrophysical relationships (Fig. 10h-i) are smoother than the case in which this uncertainty is ignored (Fig. 10g). The mean posterior fields of the petrophysical prediction uncertainty distributions (Fig. 10k-l) are very similar and correlated with the posterior velocity means.

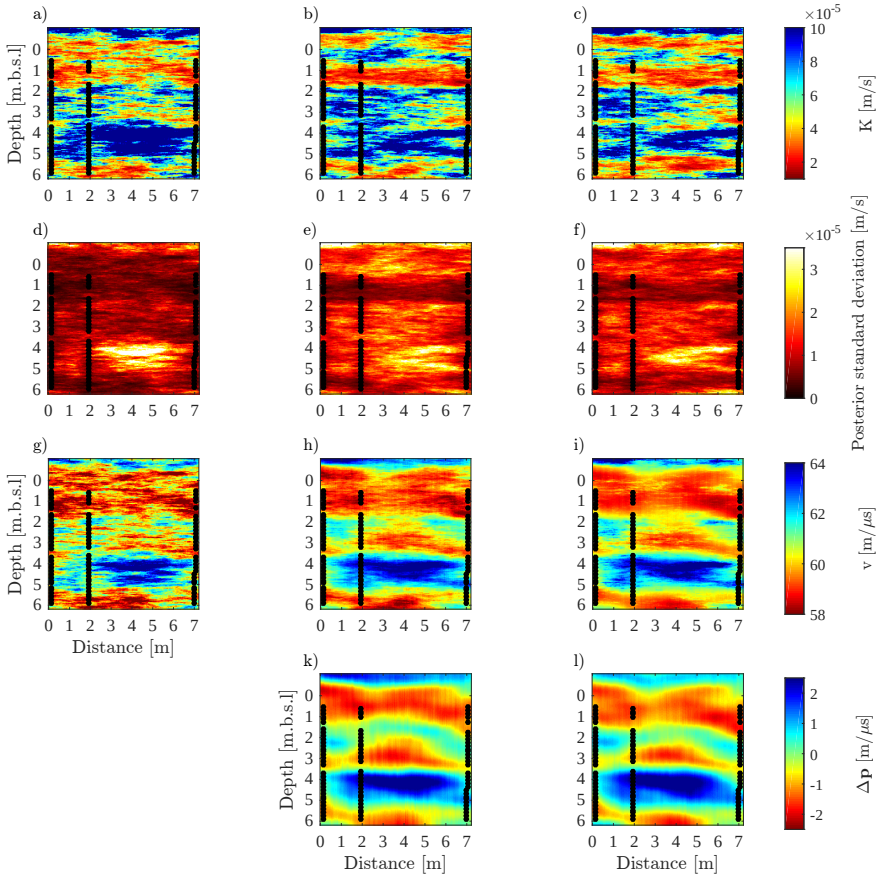


Figure 10: Mean of the posterior hydraulic conductivity, K, realizations obtained using a (a) perfect linear, (b) scattered linear and (c) scattered quadratic petrophysical relationship with the corresponding (d)-(f) standard deviation of the posterior hydraulic conductivity estimates, respectively. Mean of the posterior GPR velocity realizations obtained using (g) perfect linear, (h) scattered linear and (i) scattered quadratic petrophysical relationships. Mean of the posterior petrophysical prediction uncertainty estimates for the (k) linear and (l) quadratic petrophysical relationship. The different measurement depths of the flowmeter data (black points) are indicated for boreholes S14 (on the left), T2 (in the middle) and M3 (on the right). The posterior distributions are computed from MCMC simulation with the DREAM$_{(ZS)}$ algorithm using 8 chains with $2.5{\cdot}10^5$ iterations.

The red lines in Fig. 11a-c depict the inferred mean petrophysical relationships and the scatter (black dots) around them represents the inferred mean

petrophysical prediction uncertainty. The GPR velocity range appears to be overestimated whether $\mathbf{\Delta p}$ is ignored (Fig. 11a) or accounted for together with a quadratic petrophysical model (Fig. 11c), while a scattered linear petrophysical relationship (Fig. 11b) provides a velocity range in agreement with previous studies (Hubbard et al., 2001; Chen et al., 2001; Linde et al., 2008; Linde & Vrugt, 2013; Brunetti et al., 2017).

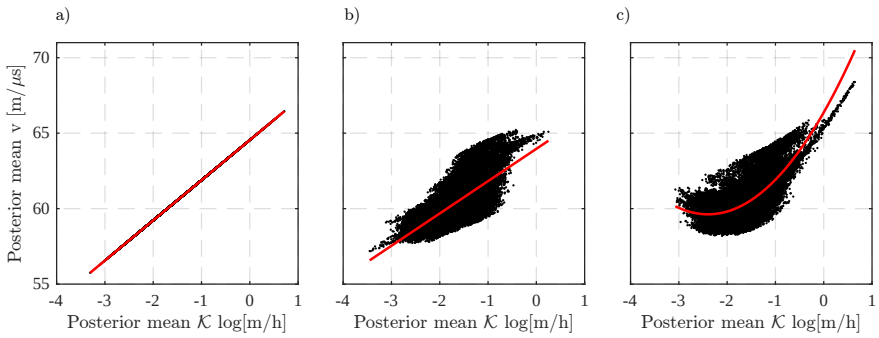


Figure 11: Scatter plots of the mean posterior hydraulic conductivity estimates against the mean posterior GPR velocity estimates assuming a (a) perfect linear, (b) scattered linear and (c) scattered quadratic petrophysical relationship. The red lines depict the inferred mean petrophysical relationship, while the scatter represents the inferred mean petrophysical prediction uncertainty.

We now turn our attention to the Bayesian model selection results. We find that Model 2 (scattered linear relationship) has the largest evidence value (-260.20 in $\log_{10}$ units) and Model 1 ($\mathbf{\Delta p}$ are ignored) has the lowest one (-361.00) (Fig. 12). The Bayes factor for the "best" petrophysical model (Model 2) with respect to Model 1 and Model 3 is $10^{100.80}$ and $10^{9.38}$, respectively. These results confirm that the perfect petrophysical model (Model 1) is erroneous. Furthermore, the results suggest that the use of a more complex petrophysical relationship is not necessarily favoured. Even if predictions based on the quadratic petrophysical model (Model 3) fits the data slightly better than the linear petrophysical model (Model 2) (Fig. 9c), the highest evidence is found for Model 2. This is a consequence of the trade-off between parsimony and goodness of fit typical of the Occam's razor principle on which Bayesian

model selection is based.

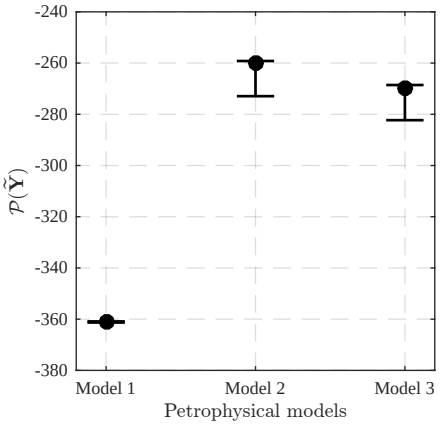


Figure 12: Mean values of the evidence in $\log_{10}$ space, $\mathcal{P}(\widetilde{\mathbf{Y}})$, and corresponding uncertainty (error bars) derived from the Gaussian mixture importance sampling method for (Model 1) a perfect linear petrophysical relationship as shown in Eq. (9), (Model 2) scattered linear petrophysical relationship such that $\mathbf{\Delta p}$ is taken into account as shown in Eqs. (9) and (11), (Model 3) scattered quadratic petrophysical relationship where $\mathbf{\Delta p}$ is taken into account as shown in Eqs. (10)-(11).

5. Discussion

Our coupled Bayesian hydrogeophysical inversion approach with explicit inference of spatially-correlated petrophysical prediction uncertainty leads to less bias (e.g., in the inferred variance of the inferred hydrogeological property field), more realistic uncertainty quantification and less over confident model selection compared to the common choice of ignoring this type of uncertainty. Even if our approach to infer petrophysical prediction uncertainty doubles the number of parameters in the inversion problem, we observe dramatic gains in sampling efficiency compared to MC-within-MCMC (e.g., Bosch (1999, 2016)). Moreover, DREAM$_{\text{(ZS)}}$ allows for parallel evaluation of the different Markov chains and, therefore, enables feasible computational times even in high (e.g., in our case, more than 200) model dimensions. Our synthetic and field-based case-studies suggest that it is not always possible to independently constrain hydrogeologi-

cal and petrophysical properties. This trade-off is particularly acute when the petrophysical prediction errors have similar geostatistical properties (e.g., orientations and integral scales) as the hydrogeological property field of interest (Fig. 7). A manifestation of this trade-off is given by the field application at the South Oyster Bacterial Transport Site, for which it was necessary to constrain the standard deviation of petrophysical prediction uncertainty and the standard deviation of the logarithm of hydraulic conductivity. Without such constraints, the inversion yields largely uncorrelated log-hydraulic conductivity and GPR velocity fields, results that are inconsistent with previous studies (Chen et al., 2001; Hubbard et al., 2001; Linde et al., 2008). This suggests that a careful petrophysical analysis involving borehole data or literature reviews are needed to define constraining prior information when performing coupled hydrogeophysical inversion of field data.

In a previous study on Bayesian hydrogeophysical inversion model selection that ignored petrophysical prediction uncertainty (Brunetti et al., 2017), it was found that the typically large data sets encountered in geophysics and the assumption of small uncorrelated data errors (Gaussian likelihood) lead to very strong confidence in the ability of geophysical data to discriminate between alternative conceptual hydrogeological models. By including spatially-correlated petrophysical prediction uncertainty, we find for a synthetic example (Fig. 5) that the magnitude of the Bayes factor of the "best" conceptual model relative to the worse one decreases by 63 orders of magnitude. Nevertheless, the comparison between Case 3 (petrophysical prediction errors ignored) and Case 4 (petrophysical prediction errors accounted for) in Fig. 5a and Table 3 still indicates high Bayes factors and a practically-speaking unique ability of geophysical data to find the most appropriate conceptual hydrogeological model among a set of candidates. In the future, one should also account for the effect of modelling errors (i.e., the discrepancy between actual physical responses and those simulated with simplified physics; here, a ray-based approximation in the present study instead of a full solution of the Maxwell's equations). A number of promising approaches to address modelling errors are available (Brynjarsdóttir

& O'Hagan, 2014; Hansen et al., 2014; Xu & Valocchi, 2015). Accounting for modelling errors is an essential next step to achieve reliable Bayesian hydrogeophysical model selection; we anticipate that this will further decrease the range of Bayes factors.

Bayesian model selection at the South Oyster Bacterial Transport Site (Section 4) demonstrates clearly that the relationship between log-hydraulic conductivity and GPR velocity is not a perfect relationship. That is, the petrophysical model with a scattered linear relationship has a much higher evidence than results obtained by assuming a perfect linear relationship. However, contrasting results were obtained in the synthetic example of Section 3.3 that did not involve any hydrogeological point measurements. In that case, formal Bayesian model selection erroneously favoured a conceptual model that ignored petrophysical prediction uncertainty. This happens because this conceptual model has fewer parameters and is still able to fit the data well, albeit with a porosity model with biased variance. At the South Oyster Bacterial transport Site, we condition all model proposals to point data (flowmeter estimates of hydraulic conductivity) and it is then not possible to propose a biased model close to the boreholes. Hence, the scattered petrophysical relationship is preferred. However, even if the inclusion of point conditioning in the synthetic example (not shown) decreased the Bayes factor, the model selection still favoured the wrong conceptual model. In the synthetic example, we considered boreholes at the left and right sides of the model domain, and the relative petrophysical prediction uncertainty was much smaller than for the field example. This could explain why the inconsistency between point data and GPR data is more evident for the field example, which led the Bayesian model selection to favour a model with petrophysical prediction uncertainty. These findings suggest that MCMC inversion and model selection is not always able to identify the "right" model and that their outputs need to be treated with some caution. The more prior information that is available (e.g., on petrophysical prediction uncertainty in terms of variance and correlation scale), the more reliable are the results. Indeed, Bayesian model selection is built on the principle of Occam's razor and

a problem-specific and conceptual-model specific level of informative data is needed to overcome this tendency to favour a simpler, but erroneous conceptual model (e.g., Schöniger et al. (2015a)).

In this study, we have made the choice to infer for petrophysical prediction uncertainty, instead of accounting for its effects in the likelihood function. For linear theory, it is indeed possible to propagate the impact of (multi-Gaussian) petrophysical errors and add the corresponding covariance matrices to the data covariance matrix (Bosch (2004); Bosch et al. (2009); Bosch (2016); Chen & Dickens (2009)). This is not possible for non-linear theory, as the resulting impact of petrophysical uncertainty on the data leads to model-dependent non-Gaussian distributions. The corresponding problem formulation and ways to address this problem was recently discussed by Linde et al. (2017) in their Section 5.2. In the future, it would be interesting to compare these two approaches (i.e., inferring for petrophysical uncertainty (this study) or accounting for the effect of petrophysical uncertainty in the likelihood function).

6. Conclusions

We have demonstrated the importance of accounting for petrophysical prediction uncertainty in coupled hydrogeophysical inversion and highlighted the critical role played by its spatial correlation. As MCMC inversions are primarily performed to enable accurate uncertainty quantification, we suggest that petrophysical prediction uncertainty should be accounted for in future hydrogeophysical studies. In this work, we parameterize the petrophysical prediction uncertainty as a multi-Gaussian field that is inferred together with hydrogeological target properties. To decrease model dimensionality, future work should also focus on developing computationally efficient and accurate approaches to account for this uncertainty in the likelihood function.

Inferring petrophysical prediction uncertainty with MCMC leads to dramatic performance gains compared to previous work, in which it has been accounted for by Monte Carlo sampling. In our examples, we show that ignoring petro-

physical prediction uncertainty and (above all) its spatial correlation causes bias in the inferred variance of the hydrogeological properties, which implies overly variable fields. Accounting for this error source allows for consistent hydrogeological estimates and widens the estimated posterior distributions. However, the geostatistical model describing petrophysical prediction uncertainty is only partially recoverable by the inversion. When performing Bayesian model selection, accounting for petrophysical prediction uncertainty reduces overconfidence in the ability of geophysical data to discriminate between conceptual hydrogeological models of the subsurface. When considering geophysical data alone, there is a risk that Bayesian hydrogeophysical model selection will favour a model parameterization that ignores petrophysical prediction uncertainty provided that the resulting overly variable hydrogeological estimates can explain the geophysical data well. This highlights the importance of including constraining prior information about petrophysical prediction uncertainty and the value of combining geophysical and hydrogeological data in the inversion.

Acknowledgements

This work was supported by the Swiss National Science Foundation under grant number 200021_155924. We thank Eric Laloy for making the code available for building the multi-Gaussian model. We thank Jasper Vrugt for making DREAM$_{(ZS)}$ available and for providing an initial version of the code used for Gaussian mixture importance sampling. We are grateful to John Peterson and Susan Hubbard, who provided the crosshole GPR and flowmeter data from the South Oyster Bacterial Transport Site. We also thank the Editor Paolo D'Odorico, the Associated Editor Harrie-Jan Hendricks-Franssen, Anneli Guthke and two anonymous reviewers for their constructive and insightful comments.

References

Bosch, M. (1999). Lithologic tomography: From plural geophysical data to lithology estimation. *J Geophys Res-Sol Ea*, *104*, 749–766. doi:10.1029/1998JB900014.

Bosch, M. (2004). The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction. *Geophysics*, *69*, 1272–1282. doi:10.1190/1.1801944.

Bosch, M. (2016). Inference Networks in Earth Models with Multiple Components and Data. In *Integrated Imaging of the Earth: Theory and Applications* chapter 3. (pp. 29–47). John Wiley & Sons, Inc. doi:10.1002/9781118929063.ch3.

Bosch, M., Carvajal, C., Rodrigues, J., Torres, A., Aldana, M., & Sierra, J. (2009). Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir. *Geophysics*, *74*, O1–O15. doi:10.1190/1.3043796.

Brunetti, C., Linde, N., & Vrugt, J. A. (2017). Bayesian model selection in hydrogeophysics: Application to conceptual subsurface models of the South Oyster Bacterial Transport Site, Virginia, USA. *Adv Water Resour*, *102*, 127–141. doi:10.1016/j.advwatres.2017.02.006.

Brynjarsdóttir, J., & O'Hagan, A. (2014). Learning about physical parameters: The importance of model discrepancy. *Inverse Probl*, *30*, 114007. doi:10.1088/0266-5611/30/11/114007.

Chen, J., & Dickens, T. A. (2009). Effects of uncertainty in rock-physics models on reservoir parameter estimation using seismic amplitude variation with angle and controlled-source electromagnetics data. *Geophys Prospect*, *57*, 61–74. doi:10.1111/j.1365-2478.2008.00721.x.

Chen, J., Hubbard, S., Peterson, J., Williams, K., Fienen, M., Jardine, P., & Watson, D. (2006). Development of a joint hydrogeophysical inversion

approach and application to a contaminated fractured aquifer. *Water Resour Res*, *42*. doi:10.1029/2005WR004694.

Chen, J., Hubbard, S., & Rubin, Y. (2001). Estimating the hydraulic conductivity at the South Oyster Site from geophysical tomographic data using Bayesian techniques based on the normal linear regression model. *Water Resour Res*, *37*, 1603–1613. doi:10.1029/2000WR900392.

Chen, J., Hubbard, S., Rubin, Y., Murray, C., Roden, E., & Majer, E. (2004). Geochemical characterization using geophysical data and Markov Chain Monte Carlo methods: A case study at the South Oyster bacterial transport site in Virginia. *Water Resour Res*, *40*, 1–14. doi:10.1029/2003WR002883.

Copty, N., Rubin, Y., & Mavko, G. (1993). Geophysical-hydrological identification of field permeabilities through Bayesian updating. *Water Resour Res*, *29*, 2813–2825. doi:10.1029/93WR00745.

Day-Lewis, F. D., Singha, K., & Binley, A. M. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *J Geophys Res-Sol Ea*, *110*, 1–17. doi:10.1029/2004JB003569.

Doetsch, J., Linde, N., Coscia, I., Greenhalgh, S. A., & Green, A. G. (2010). Zonation for 3D aquifer characterization based on joint inversions of multimethod crosshole geophysical data. *Geophysics*, *75*, G53–G64. doi:10.1190/1.3496476.

Doyen, P. M. (1988). Porosity from seismic data: A geostatistical approach. *Geophysics*, *53*, 1263–1275. doi:10.1190/1.1442404.

Doyen, P. M. (2007). Seismic reservoir characterization: An Earth Modelling Perspective. *EAGE publications*, (p. 255). doi:10.3997/9789073781771.

Gallardo, L. A., & Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophys Res Letters*, *30*. doi:10.1029/2003GL017370.

Gelman, A., Roberts, G. O., Gilks, W. R. et al. (1996). Efficient Metropolis jumping rules. *Bayesian statistics*, *5*, 599–608.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat Sci*, *7*, 457–472. doi:10.1214/ss/1177011136.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain Monte Carlo in practice*. CRC press.

Grana, D., & Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, *75*, O21–O37. doi:10.1190/1.3386676.

Grana, D., Pirrone, M., & Mukerji, T. (2012). Quantitative log interpretation and uncertainty propagation of petrophysical properties and facies classification from rock-physics modeling and formation evaluation analysis. *Geophysics*, *77*, WA45–WA63. doi:10.1190/geo2011-0272.1.

Gull, S. F. (1988). Bayesian inductive inference and maximum entropy. In *Maximum-entropy and Bayesian methods in Science and Engineering* (pp. 53–74). Springer volume 31-32. doi:10.1007/978-94-009-3049-0_4.

Haber, E., & Oldenburg, D. (1997). Joint inversion: A structural approach. *Inverse Probl*, *13*, 63. doi:10.1088/0266-5611/13/1/006.

Hammersley, J. M., & Handscomb, D. C. (1964). *Monte Carlo methods* volume 1. Springer Netherlands. doi:10.1007/978-94-009-5819-7.

Hansen, T., Cordua, K., Jacobsen, B., & Mosegaard, K. (2014). Accounting for imperfect forward modeling in geophysical inverse problems - Exemplified for crosshole tomography. *Geophysics*, *79*, H1–H21. doi:10.1190/geo2013-0215.1.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*, 97–109. doi:10.1093/biomet/57.1.97.

Hermans, T., Oware, E., & Caers, J. (2016). Direct prediction of spatially and temporally varying physical properties from time-lapse electrical resistance data. *Water Resour Res*, *52*, 7262–7283. doi:10.1002/2016WR019126.

Hinnell, A., Ferré, T., Vrugt, J., Huisman, J., Moysey, S., Rings, J., & Kowalsky, M. (2010). Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. *Water Resour Res*, *46*. doi:10.1029/2008WR007060.

Hubbard, S. S., Chen, J., Peterson, J., Majer, E. L., Williams, K. H., Swift, D. J., Mailloux, B., & Rubin, Y. (2001). Hydrogeological characterization of the South Oyster Bacterial Transport Site using geophysical data. *Water Resour Res*, *37*, 2431–2456. doi:10.1029/2001WR000279.

Jefferys, W. H., & Berger, J. (1992). Ockham's Razor and Bayesian Analysis. *Am Sci*, *80*, 64–72.

Jeffreys, H. (1935). Some Tests of Significance, Treated by the Theory of Probability. *Math Proc Cambridge*, *31*, 203–222. doi:10.1017/S030500410001330X.

Jeffreys, H. (1939). *Theory of Probability*. (3rd ed.). Oxford University Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *J Am Stat Assoc*, *90*, 773–795. doi:10.1080/01621459.1995.10476572.

Kowalsky, M. B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., & Gee, G. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resour Res*, *41*. doi:10.1029/2005WR004237.

Laloy, E., Linde, N., Jacques, D., & Vrugt, J. A. (2015). Probabilistic inference of multi-Gaussian fields from indirect hydrological data using circulant embedding and dimensionality reduction. *Water Resour Res*, *51*, 4224–4243. doi:10.1002/2014WR016395.

Laloy, E., & Vrugt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try DREAM_{ZS} and high-performance computing. *Water Resour Res*, *48*, 1–18. doi:10.1029/2011WR010608.

Linde, N. (2014). Falsification and corroboration of conceptual hydrological models using geophysical data. *Wiley Interdiscip Rev Water*, *1*, 151–171. doi:10.1002/wat2.1011.

Linde, N., Binley, A., Tryggvason, A., Pedersen, L. B., & Revil, A. (2006). Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data. *Water Resour Res*, *42*. doi:10.1029/2006WR005131.

Linde, N., Ginsbourger, D., Irving, J., Nobile, F., & Doucet, A. (2017). On Uncertainty Quantification in Hydrogeology and Hydrogeophysics. *Adv Water Resour*, *110*, 166–181. doi:10.1016/j.advwatres.2017.10.014.

Linde, N., Tryggvason, A., Peterson, J. E., & Hubbard, S. S. (2008). Joint inversion of crosshole radar and seismic traveltimes acquired at the South Oyster Bacterial Transport Site. *Geophysics*, *73*, G29–G37. doi:10.1190/1.2937467.

Linde, N., & Vrugt, J. A. (2013). Distributed soil moisture from crosshole ground-penetrating radar travel times using stochastic inversion. *Vadose Zone J*, *12*, 1–21. doi:10.2136/vzj2012.0101.

Lochbühler, T., Breen, S. J., Detwiler, R. L., Vrugt, J. A., & Linde, N. (2014). Probabilistic electrical resistivity tomography of a CO_2 sequestration analog. *J Appl Geophys*, *107*, 80–92. doi:10.1016/j.jappgeo.2014.05.013.

Lochbühler, T., Doetsch, J., Brauchler, R., & Linde, N. (2013). Structure-coupled joint inversion of geophysical and hydrological data. *Geophysics*, *78*, ID1–ID14. doi:10.1190/geo2012-0460.1.

Lochbühler, T., Vrugt, J. A., Sadegh, M., & Linde, N. (2015). Summary statistics from training images as prior information in probabilistic inversion. *Geophys J Int*, *201*, 157–171. doi:10.1093/gji/ggv008.

MacKay, D. J. (1992). Bayesian interpolation. *Neural Comput*, *4*, 415–447. doi:10.1162/neco.1992.4.3.415.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *J Chem Phys*, *21*, 1087–1092. doi:10.1063/1.1699114.

Mukerji, T., Jørstad, A., Avseth, P., Mavko, G., & Granli, J. (2001). Mapping lithofacies and pore-fluid probabilities in a north sea reservoir: Seismic inversions and statistical rock physics. *Geophysics*, *66*, 988–1001. doi:10.1190/1.1487078.

Oware, E., Moysey, S., & Khan, T. (2013). Physically based regularization of hydrogeophysical inverse problems for improved imaging of process-driven systems. *Water Resour Res*, *49*, 6238–6247. doi:10.1002/wrcr.20462.

Podvin, P., & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophys J Int*, *105*, 271–284. doi:10.1111/j.1365-246X.1991.tb03461.x.

Pride, S. (1994). Governing equations for the coupled electromagnetics and acoustics of porous media. *Phys Rev B*, *50*, 15678–15696. doi:10.1103/PhysRevB.50.15678.

Robert, C., & Casella, G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media. doi:10.1007/978-1-4757-4145-2.

Roth, K., Schulin, R., Fluhler, H., & Attinger, W. (1990). Using a composite dielectric approach. *Water Resour Res*, *26*, 2267–2273. doi:10.1029/WR026i010p02267.

947 Rubin, Y., Mavko, G., & Harris, J. (1992). Mapping permeability in hetero-
948 geneous aquifers using hydrologic and seismic data. *Water Resour Res*, *28*,
949 1809–1816. doi:10.1029/92WR00154.

950 Scheibe, T. D., Hubbard, S. S., Onstott, T. C., & DeFlaun, M. F. (2011).
951 Lessons learned from Bacterial Transport Research at the South Oyster Site.
952 *Groundwater*, *49*, 745–763. doi:10.1111/j.1745-6584.2011.00831.x.

953 Schöniger, A., Illman, W. A., Wöhling, T., & Nowak, W. (2015a). Finding
954 the right balance between groundwater model complexity and experimental
955 effort via Bayesian model selection. *J Hydrol*, *531*, 96–110. doi:10.1016/j.
956 jhydrol.2015.07.047.

957 Schöniger, A., Wöhling, T., & Nowak, W. (2015b). A statistical concept to
958 assess the uncertainty in Bayesian model weights and its impact on model
959 ranking. *Water Resour Res*, *51*, 7524–7546. doi:10.1002/2015WR016918.

960 Shahraeeni, M. S., & Curtis, A. (2011). Fast probabilistic nonlinear petrophys-
961 ical inversion. *Geophysics*, *76*, E45–E58. doi:10.1190/1.3540628.

962 Shahraeeni, M. S., Curtis, A., & Chao, G. (2012). Fast probabilistic petrophys-
963 ical mapping of reservoirs from 3D seismic data. *Geophysics*, *77*, O1–O19.
964 doi:10.1190/geo2011-0340.1.

965 Sun, J., & Li, Y. (2016). Joint inversion of multiple geophysical data using
966 guided fuzzy c-means clustering. *Geophysics*, *81*, ID37–ID57. doi:10.1190/
967 geo2015-0457.1.

968 Sun, J., & Li, Y. (2017). Joint inversion of multiple geophysical and petrophys-
969 ical data using generalized fuzzy clustering algorithms. *Geophys J Int*, *208*,
970 1201–1216. doi:10.1093/gji/ggw442.

971 Vasco, D., Daley, T. M., & Bakulin, A. (2014). Utilizing the onset of time-
972 lapse changes: A robust basis for reservoir monitoring and characterization.
973 *Geophys J Int*, *197*, 542–556. doi:10.1093/gji/ggt526.

974 Volpi, E., Schoups, G., Firmani, G., & Vrugt, J. (2017). Sworn Testimony of the
975 Model Evidence: Gaussian Mixture Importance (GAME) sampling. *Water*
976 *Resour Res*, . doi:10.1002/2016WR020167.

977 Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM
978 software package: Theory, concepts, and MATLAB implementation. *Environ*
979 *Modell Softw*, *75*, 273–316. doi:10.1016/j.envsoft.2015.08.013.

980 Xu, T., & Valocchi, A. (2015). A Bayesian approach to improved calibration
981 and prediction of groundwater models with structural error. *Water Resour*
982 *Res*, *51*, 9290–9311. doi:10.1002/2015WR017912.